White Matter Hyperintensities Quantification in Healthy Adults: A Systematic Review and Meta-Analysis

Luca Melazzini, MD,[1] Paolo Vitali, MD, PhD,[2] Emanuele Olivieri, MD,[3] Marco Bolchini, MD,[4] Moreno Zanardo, MSc, PhD,[2] Filippo Savoldi, MD,[5] Giovanni Di Leo, MSc,[2] Ludovica Griffanti, MSc, PhD,[6] Giuseppe Baselli, MSc,[7] Francesco Sardanelli, MD,[1,2*] and Marina Codari, MSc, PhD[8]

Background: Although white matter hyperintensities (WMH) volumetric assessment is now customary in research studies, inconsistent WMH measures among homogenous populations may prevent the clinical usability of this biomarker.
Purpose: To determine whether a point estimate and reference standard for WMH volume in the healthy aging population could be determined.
Study Type: Systematic review and meta-analysis.
Population: In all, 9716 adult subjects from 38 studies reporting WMH volume were retrieved following a systematic search on EMBASE.
Field Strength/Sequence: 1.0T, 1.5T, or 3.0T/fluid-attenuated inversion recovery (FLAIR) and/or proton density/T_2-weighted fast spin echo sequences or gradient echo T_1-weighted sequences.
Assessment: After a literature search, sample size, demographics, magnetic field strength, MRI sequences, level of automation in WMH assessment, study population, and WMH volume were extracted.
Statistical Tests: The pooled WMH volume with 95% confidence interval (CI) was calculated using the random-effect model. The I^2 statistic was calculated as a measure of heterogeneity across studies. Meta-regression analysis of WMH volume on age was performed.
Results: Of the 38 studies analyzed, 17 reported WMH volume as the mean and standard deviation (SD) and were included in the meta-analysis. Mean and SD of age was 66.11 ± 10.92 years (percentage of men 50.45% ± 21.48%). Heterogeneity was very high (I^2 = 99%). The pooled WMH volume was 4.70 cm^3 (95% CI: 3.88–5.53 cm^3). At meta-regression analysis, WMH volume was positively associated with subjects' age (β = 0.358 cm^3 per year, $P < 0.05$, $R^2 = 0.27$).
Data Conclusion: The lack of standardization in the definition of WMH together with the high technical variability in assessment may explain a large component of the observed heterogeneity. Currently, volumes of WMH in healthy subjects are not comparable between studies and an estimate and reference interval could not be determined.
Level of Evidence: 1
Technical Efficacy Stage: 1

J. MAGN. RESON. IMAGING 2021;53:1732–1743.

WHITE MATTER HYPERINTENSITIES (WMH) seen on T_2-weighted magnetic resonance imaging (MRI) brain scans are common radiological findings in adults, associated with a higher risk for developing stroke, dementia, gait disturbances, and psychiatric diseases.[1,2] Together with lacunes, microbleeds, and enlarged perivascular

View this article online at wileyonlinelibrary.com. DOI: 10.1002/jmri.27479

Received Sep 24, 2020, Accepted for publication Dec 3, 2020.

*Address reprint requests to: F.S., Department of Biomedical Sciences for Health, Università degli Studi di Milano, Unit of Radiology, IRCCS Policlinico San Donato, Via Morandi 30, 20097 San Donato Milanese, Milan, Italy. E-mail: francesco.sardanelli@unimi.it

From the [1]Department of Biomedical Sciences for Health, Università degli Studi di Milano, Milan, Italy; [2]Unit of Radiology, IRCCS Policlinico San Donato, San Donato Milanese, Italy; [3]Medicine and Surgery Medical School, Università degli Studi di Milano, Milano, Italy; [4]Department of Clinical and Experimental Sciences, Università degli Studi di Brescia, Brescia, Italy; [5]Postgraduate School in Radiology, Università degli Studi di Milano, Milano, Italy; [6]Department of Psychiatry, Wellcome Centre for Integrative Neuroimaging (WIN), University of Oxford, Oxford, UK; [7]Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milano, Italy; and [8]Department of Radiology, Stanford University School of Medicine, Stanford, California, USA

Additional supporting information may be found in the online version of this article

spaces, WMH are considered a neuroimaging biomarker of cerebral small vessel disease (SVD).[3]

The appearance of WMH is heterogeneous, ranging from focal to diffuse confluent lesions.[4] Heterogeneity may result from a combination of different underlying etiology and histopathological changes.[5] In the general population, WMH volume primarily depends on age more than other vascular risk factors.[6] Modeling WMH volume heterogeneity in the healthy aging population would therefore be important for understanding its use in brain disease.[7]

Volumetry of WMH has become customary in research due to the availability of methods with high sensitivity for detecting small lesions and better reliability when compared to qualitative or semiquantitative methods.[8] Nevertheless, differences in imaging sequences, protocol parameters, and display settings can result in poor reproducibility of manual quantification when dealing with small WMH volumes.[9] In this respect, the use of automated methods with intensity normalization has been strongly advised to improve reproducibility.[10]

Variability in quantitative assessment may arise at any stage, from image acquisition to image postprocessing and interpretation of radiological findings.[11] The development and validation of a strong neuroimaging biomarker passes through different stages, from single-center cross-sectional to large-scale studies.[12] As far as WMH are concerned, some aspects of validation are lacking, with limited comparisons among segmentation tools and few longitudinal studies.[13]

The reproducibility estimation of WMH is crucial for the development of a surrogate biomarker for cerebral SVD. Bias sources in WMH volume estimation should be controlled as much as possible, enabling comparison among different studies.[11] Studies on young subjects with small lesion volumes should also warrant high accuracy and good reproducibility. The same applies when following up subjects for longitudinal WMH volume changes.[8]

Thus, the purpose of this study was to investigate the current source of bias in WMH volume assessment in healthy adults by conducting a systematic review and meta-analysis on currently available studies. By analyzing the variability in image acquisition protocols, postprocessing automatization levels, and reported WHM volumes, we aimed to investigate whether a point estimate and reference standard for WMH volume could be determined.

Materials and Methods

No Ethics Committee approval was needed for this systematic review. The study protocol was registered on PROSPERO (BLINDED, available from: https://www.crd.york.ac.uk/prospero/) and reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses.[14]

Search Strategy and Eligibility Criteria

A systematic search of the literature was performed on 8 January 2020 using EMBASE (Elsevier, Amsterdam, Netherlands) for articles focusing on WMH quantification in healthy adults. A controlled vocabulary (EMBASE thesaurus) was used.

The exact search query was: ("white matter hyperintensity"/exp OR "white matter hyperintensity" OR "leukoaraiosis" OR "wmh" OR "wml") AND ("quantitative analysis" OR "quantification" OR "volume" OR "volumetric" OR "white matter hyperintensity volume" OR "white matter hyperintensity volume"/exp) AND ("aging" OR "aging"/exp OR "elderly" OR "aged" OR "aged"/exp OR "elderly"/exp) AND ("normal human"/exp OR "normal human" OR "healthy").

The search was limited to original studies on humans, published since 2008 in peer-reviewed journals, written in English and provided with an abstract.

Three independent readers performed initial screening based on the title and abstract only (L.M., 4 years of experience in medical imaging; M.B. and E.O., 1 year of experience each). Eligible articles were those that reported in the abstract that quantification of WMH volume in healthy participants was performed. Subjects from community-dwelling longitudinal cohorts or control subjects from case–control studies were deemed healthy unless any neurological, psychiatric, and cognitive conditions were explicitly stated. When provided, detailed information about subjects' health status was further characterized as the absence of neurological, psychiatric, cognitive, cardiovascular or major (both inherited and acquired) diseases.

Eligible articles were then retrieved and read in full by the same three readers who performed the initial screening. Disagreements between readers were resolved by the decision of a fourth independent reviewer (M.C., 7 years of experience in medical imaging), who had the final say over any disagreement between the three readers.

Data Extraction

Data extraction was independently performed by the same reviewers who performed the initial screening using the same two-step selection protocol. Articles with unclear presentation of data (eg, unspecified unit of measurement for WMH volumes) or missing WMH volumes data were discarded. Finally, in order to perform the subsequent meta-analysis, only articles that reported the total WMH volume expressed as mean and standard deviation (SD) were included.

Data extracted included: first author's family name and year of publication; subjects' demographics; sample size and clinical history and WMH volume of participants included in each study. Moreover, we collected the level of automation for WMH volume assessment (manual, semiautomated, or automated); magnetic field strength; MRI sequences adopted.

Statistical Analysis

Statistical analysis was performed using Comprehensive Meta-Analysis v. 2.2.057 (Biostat, Englewood, NJ). First, the I^2 statistic was calculated, which estimates the percentage of variability across studies that is due to heterogeneity rather than chance.[15,16] The random-effect model with the DerSimonian and Laird method, suitable for handling heterogeneous data, was used to calculate the pooled WMH volume and its 95% confidence interval (CI).[17] If statistical heterogeneity was low, a pooled SD would be calculated as the root mean square of all contributing study SD values to take into account each study sample size, and the reference range built as the WMH pooled volume ± 2 pooled SD.[18]

Potential sources of heterogeneity were evaluated by meta-regression and subgroup analysis.

Meta-regression analysis was performed to assess the correlation between WMH volume and subjects' age. Due to nonlinearity of WMH changes with age,[19] we performed a second meta-regression analysis after excluding all studies involving young subjects (<50 years old). This threshold was established as subjects less than 50 years old are expected to carry an extremely low WMH burden.[6]

Subgroup analysis with the random-effect model was performed to investigate the effect of the magnetic field strength and the level of automation used for WMH assessment on WMH volume.

The statistical significance threshold was set at $P < 0.05$.

Risk of Bias

For quality appraisal, we relied on the risk of bias assessment using RevMan v. 5.4.1 (Review Manager software, Cochrane Collaboration) as performed in a similar study.[20] We selected the risk of bias items excluding those referred to randomization ("random sequence generation") and blinding ("blinding of participants" and "personnel and blinding of outcome assessment"), as they were unrelated to the design of this review. We evaluated the "allocation concealment" (selection bias) by assessing the characterization of the health status of subjects included in the analyzed studies. "Incomplete outcome data" (attrition bias) referred to the handling of potential missing data, while "selective reporting" (reporting bias) was suspected if the analyzed data were not fully reported. For the "other bias" item we considered potential confounders related to WMH volume assessment (imaging protocols and operators' years of experience when manual segmentation was performed).

The risk of publication bias was instead assessed by visually inspecting funnel plots and performing the Egger test and Kendall's tau.[16]

Results

Retrieved and Selected Literature

A flowchart of the literature search is shown in Fig. 1. From the initial search, 162 articles were retrieved with 38 of them being analyzed.[19,21–57] All included studies were prospective;

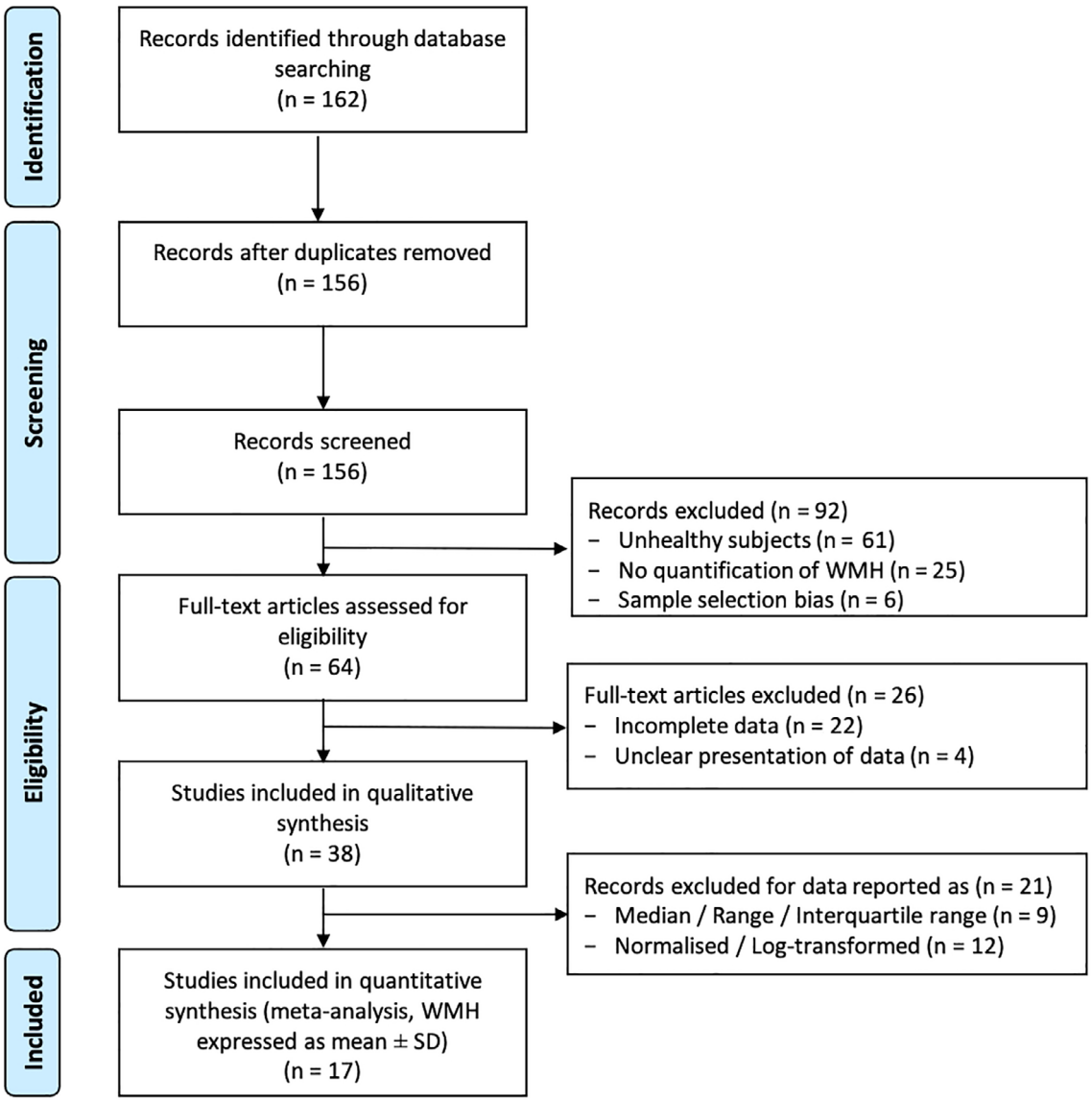

FIGURE 1: Flow chart of the literature search. Of the 162 initially retrieved articles, 38 were included in the qualitative synthesis and 17 in the meta-analysis.

TABLE 1. Main Characteristics of the 21 Study Groups Included in the Quantitative Synthesis

Study	*N*	Men (%)	Age (years)	MFS	Sequence	Level of automation	Healthy subjects: free of	Part of population study	WMH volume (cm^3)[a]
Jäncke et al. 2019[21] *(First sex subgroup)*	107	100	71	3.0	T_1-weighted	Automated	☑Neurological ☑Psychiatric ☑Cognitive ☐Cardiovascular ☐Major diseases	N.A.	3.30 ± 3.10
Jäncke et al. 2019[21] *(second sex subgroup)*	109	0	70	3.0	T_1-weighted	Automated	☑Neurological ☑Psychiatric ☑Cognitive ☐Cardiovascular ☐Major diseases	N.A.	3.60 ± 5.10
Yatawara et al. 2019[43]	79	41	63	3.0	FLAIR	Automated	☑Neurological ☑Psychiatric ☐Cognitive ☐Cardiovascular ☐Major diseases	N.A.	4.96 ± 6.53
de Schipper et al. 2019[52]	218	63	65	3.0	FLAIR	Automated	☑Neurological ☑Psychiatric ☐Cognitive ☐Cardiovascular ☐Major diseases	Leiden Longevity Study	4.80 ± 5.10
Ye et al. 2019[57]	33	48	62	3.0	FLAIR	Semiautomated	☑Neurological ☐Psychiatric ☐Cognitive ☐Cardiovascular ☐Major diseases	N.A.	1.28 ± 1.12
Lampe et al. 2019[23]	702	54	68	3.0	FLAIR	Automated	☑Neurological ☑Psychiatric ☐Cognitive ☐Cardiovascular ☐Major diseases	LIFE (Leipzig Research Centre for Civilization Diseases)-Adult-Study	4.25 ± 6.66

TABLE 1. Continued

Study	*N*	Men (%)	Age (years)	MFS	Sequence	Level of automation	Healthy subjects: free of	Part of population study	WMH volume (cm^3)[a]
Nylander et al. 2018[26]	396	N.A.	75	1.5	PD/T_2-weighted	Semiautomated	☑Neurological ☐Psychiatric ☑Cognitive ☐Cardiovascular ☐Major diseases	Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS)	10.50 ± 5.20
Huang et al. 2018[19] *(First age subgroup)*	102	53	26	3.0	FLAIR	Automated	☑Neurological ☑Psychiatric ☑Cognitive ☑Cardiovascular ☐Major diseases	N.A.	0.11 ± 0.26
Huang et al. 2018[19] (*Second age subgroup)*	89	33	50	3.0	FLAIR	Automated	☑Neurological ☑Psychiatric ☑Cognitive ☑Cardiovascular ☐Major diseases	N.A.	0.27 ± 0.51
Huang et al. 2018[19] *(Third age subgroup)*	121	61	72	3.0	FLAIR	Automated	☑Neurological ☑Psychiatric ☑Cognitive ☑Cardiovascular ☐Major diseases	N.A.	8.17 ± 12.61
Phyu et al. 2018[27]	18	50	37	3.0	PD/T_2-weighted	Semiautomated	☑Neurological ☐Psychiatric ☐Cognitive ☐Cardiovascular ☐Major diseases	N.A.	0.20 ± 0.48
Van Rooden et al. 2018[28]	42	40	68	3.0	FLAIR	Automated	☑Neurological ☑Psychiatric ☑Cognitive ☐Cardiovascular ☐Major diseases	N.A.	7.90 ± 4.60

TABLE 1. Continued

Study	*N*	Men (%)	Age (years)	MFS	Sequence	Level of automation	Healthy subjects: free of	Part of population study	WMH volume (cm^3)[a]
Ye et al. 2017[31]	64	45	70	3.0	FLAIR	Automated	☑Neurological ☑Psychiatric ☑Cognitive ☐Cardiovascular ☑Major diseases	N.A.	3.08 ± 4.67
Ye et al. 2017[31] *(Not-cognitively declining subgroup of the sample above at 35 months follow up)*	43	44	72	3.0	FLAIR	Automated	☑Neurological ☑Psychiatric ☑Cognitive ☐Cardiovascular ☑Major diseases	N.A.	5.09 ± 7.14
De Marco et al. 2017[33]	51	31	62	3.0	FLAIR	Automated	☑Neurological ☐Psychiatric ☑Cognitive ☐Cardiovascular ☐Major diseases	N.A.	3.26 ± 5.83
Jiang et al. 2015[37]	155	43	79	3.0	FLAIR	Automated	☑Neurological ☑Psychiatric ☑Cognitive ☐Cardiovascular ☑Major diseases	Sidney Memory and Aging Study	14.16 ± 10.55
Ternifi et al 2014[39]	9	0	71	1.5	FLAIR	Semiautomated	☑Neurological ☐Psychiatric ☑Cognitive ☑Cardiovascular ☑Major diseases	N.A.	10.24 ± 6.42
Bijanki et al. 2013[42]	22	45	69	1.5	FLAIR	Manual	☑Neurological ☑Psychiatric ☑Cognitive ☑Cardiovascular ☑Major diseases	Aging, Vascular Disease, and Cognition	6.30 ± 4.50

TABLE 1. Continued

Study	*N*	Men (%)	Age (years)	MFS	Sequence	Level of automation	Healthy subjects: free of	Part of population study	WMH volume (cm^3)[a]
Makedonov et al. 2013[46]	50	46	70	1.5	PD/T_2 weighted	Semiautomated	☑Neurological ☐Psychiatric ☑Cognitive ☐Cardiovascular ☐Major diseases	N.A.	5.20 ± 6.88
Portet et al. 2012[48]	274	41	71	N.A.	T_2-weighted	Semiautomated	☑Neurological ☐Psychiatric ☑Cognitive ☑Cardiovascular ☐Major diseases	ESPRIT (Enquête de Santé Psychologique - Risques, Incidence et Traitement)	1.79 ± 3.19
Silbert et al. 2009[49]	49	47	84	1.5	PD/T_2-weighted	Manual	☑Neurological ☑Psychiatric ☑Cognitive ☑Cardiovascular ☐Major diseases	Oregon Brain Aging Study	7.50 ± 7.20

Age values are presented as the means.
MFS = magnetic field strength, FLAIR = fluid-attenuated inversion recovery, PD = proton density, WMH = white matter hyperintensities.
[a]Data are presented as mean ± standard deviation.

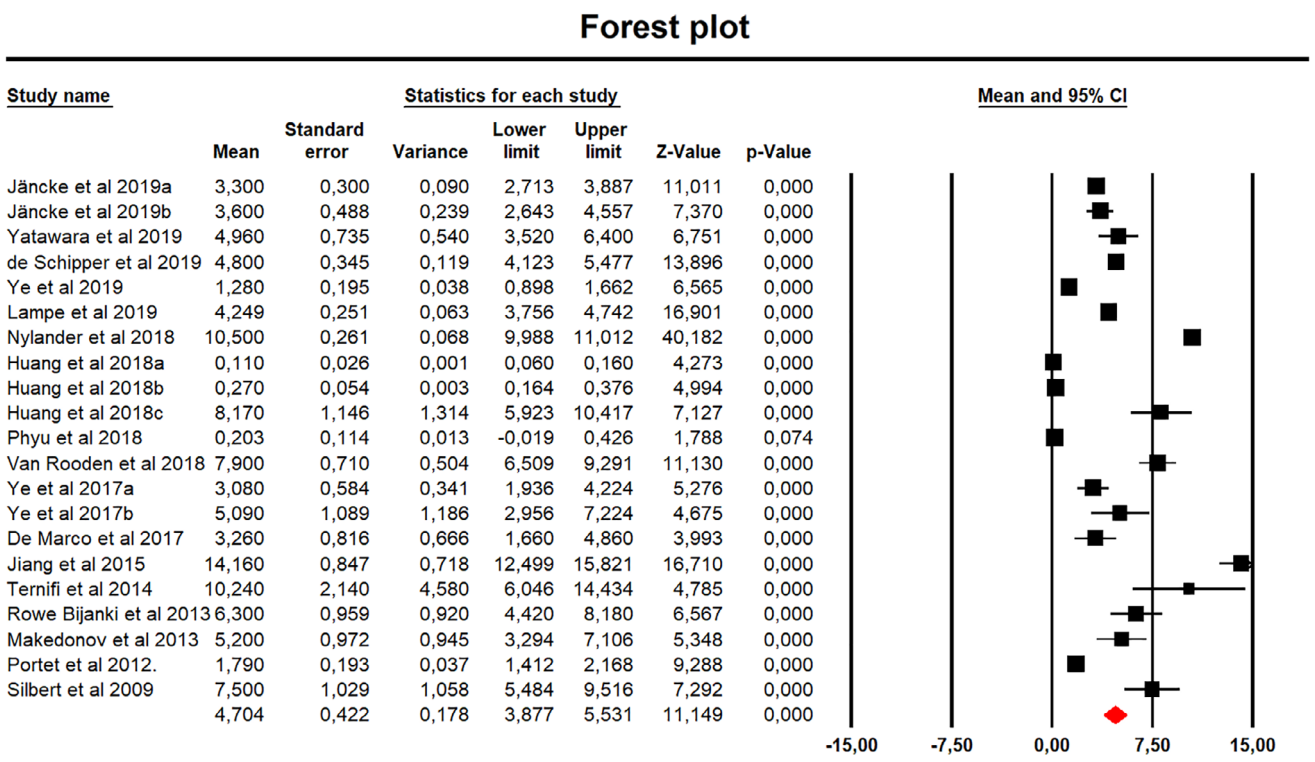

FIGURE 2: forest plot of the 17 analyzed studies, for a total of 21 independent study groups. The forest plot is a graphical representation of the results of the meta-analysis. The studies are represented by squares whose area is proportional to the weight of that study in the analysis. Since this analysis is based on the random-effects model, the weight is the inverse of the total variance (within- and between-study variance) for each study. Heterogeneity among studies was very high (I^2 = 99%). The last row shows the pooled WMH volume.

five had an intraindividual study design with two[21,24,31,40] or three study groups[19] that were examined separately. In those cases, study groups were considered to be independent, resulting in a total of 44 study groups from the 38 articles.

Qualitative Analysis of the Included Studies

The number of healthy subjects in each study group included ranged from 9[39] to 2640,[54] resulting in a total of 9716 subjects.

Participants mean and SD age was 66.11 ± 10.92 years. Mean age in the analyzed subjects ranged from 26[19] to 84 years.[49] Mean ± SD percentage of men in the reviewed articles was 50.45% ± 21.48%. None of the healthy subjects had any disclosed disease. When explicitly stated, patients' history included absence of prior neurological disorders (42/44 study groups, 95%), psychiatric illnesses (31/44 study groups, 70%), cognitive impairments (35/44 study groups, 80%), cardiovascular diseases (12/44 study groups, 27%), or major diseases (12/44 study groups, 27%).

Magnetic field strength was 1.0T in 1/44 (2%) of the contributing study groups, 1.5T in 11/44 (25%), 3T in 28/44 (64%), mixed in 2/44 (4%), and not reported in 2/44 (4%).

Out of 44 study groups, 35 (79%) used fluid-attenuated inversion recovery (FLAIR) images for WMH segmentation, six (14%) used combined proton density and T_2-weighted images, one (2%) used T_2-weighted images only, and two (4%) used T_1-weighted images only. The quantitative assessment of WMH volume was automated in 17/44 (39%) study groups, semiautomated in 25/44 (57%), and manual in 2/44 (4%). Detailed demographics of subjects included in all study groups are presented in Supplemental Table S1.

WMH Volume Meta-Analysis

Seventeen articles with a total of 21 study groups reported quantitative data for WMH volume as mean ± SD and were therefore included in the meta-analysis.[19,21,23,26–28,31,33,37,39,42,43,46,48,49,52,57] [19,21,23,26–28,31,33,37,39,42,43,46,48,49,52,57] The total number of healthy subjects included in the meta-analysis was 2743. Detailed demographics of these subjects are presented in Table 1.

In analyzed studies, WMH volume ranged from 0.11[19] (study group *a*, see Fig. 2) to 14.16 cm^3.[37] Analyzed data showed high heterogeneity (Q = 2914, degrees of freedom = 20, I^2 = 99%, $P < 0.05$). Standard deviation of WMH volumes ranged from 0.26[19] (study group *a*, see

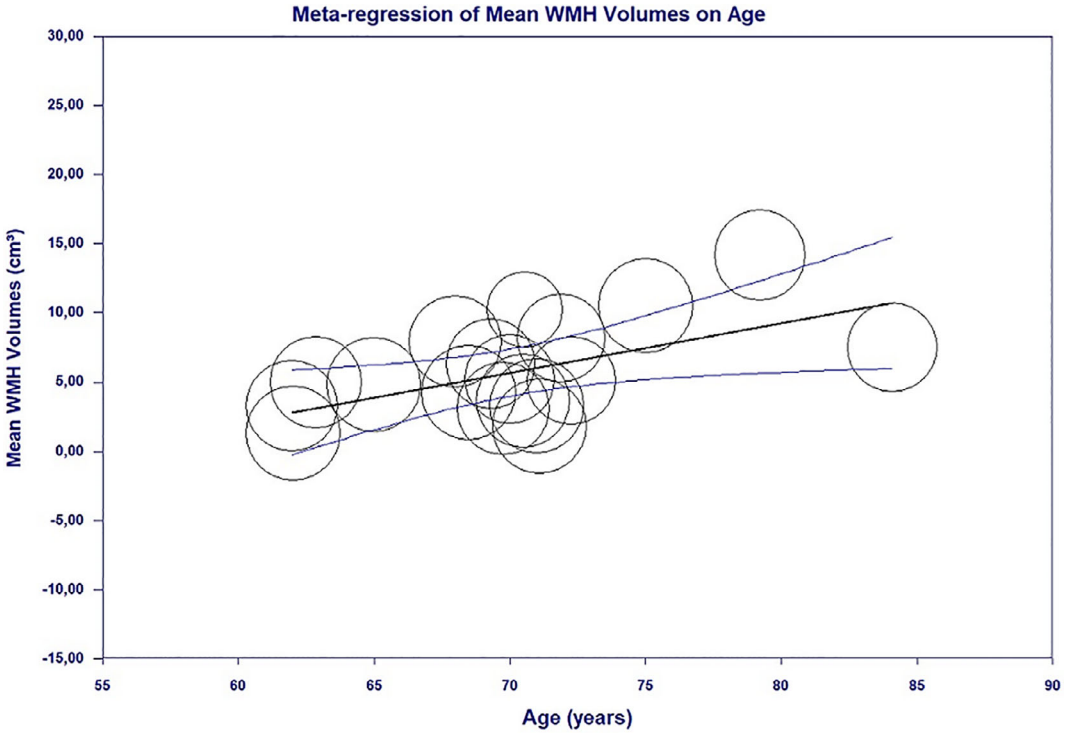

FIGURE 3: Meta-regression plot of WMH volumes on age after exclusion of three study groups with mean subjects' age <50 years and mean WMH volumes close to 0 cm^3. Eighteen study groups were included in the analysis. The area of each circle is proportional to the weight of that study in the analysis. Since this analysis was based on the random-effects model, the weight is the inverse of the total variance (within- and between-study variance) for each study.

Fig. 2) to 12.61 cm^3,[19] (study group *c*, see Fig. 2). Using the random-effect model we obtained a pooled WMH volume of 4.70 cm^3 (95% CI: 3.88–5.53 cm^3), as depicted in the forest plot (Fig. 2). The risk of bias graph (Supplemental Fig. S1) and the risk of bias summary (Supplemental Fig. S2) showed overall moderate selection and other biases and low attrition and reporting biases. At visual inspection, the funnel plot (Supplemental Fig. S3) showed a moderate risk of publication bias, as confirmed by the Egger test ($P < 0.05$) and the Kendall's tau (b = 0.033; $P > 0.05$).

Meta-Regression Analysis

Meta-regression analysis performed in the 21 study groups showed a positive statistically significant correlation between WMH volume and age (intercept β_0 = –7.030, $P < 0.05$; β_{WMH} = 0.182 cm^3 per year of age, $P < 0.05$; R^2 = 0.00) (Supplemental Fig. S4). The second meta-regression performed excluding three study groups with mean subjects' age below 50 years showed a higher positive correlation between WMH volume and subjects' age (intercept β_0 = –19.349, $P < 0.05$; β_{WMH} = 0.358 cm^3 per year of age, $P < 0.05$; R^2 = 0.27) (Fig. 3).

Subgroup Analysis

Subgroup analysis showed a significant effect on WMH volume of the magnetic field strength ($P < 0.05$). The pooled WMH volume from studies using 1.5T magnetic field strength was 7.87 cm^3 (95% CI: 5.26–10.47 cm^3). The pooled WMH volume from studies using 3.0T magnetic field strength was 3.82 cm^3 (95% CI: 3.16–4.49 cm^3) (Supplemental Fig. S5). Median age and interquartile range (IQR) for participants from studies included in the meta-analysis at 1.5T and 3.0T were 70.56 (69.65–79.55) years and 68.00 (61.98–71.00) years, respectively.

A significant difference in pooled WMH volume based on the level of automation was also found ($P < 0.05$). The pooled WMH volume from studies using a manual method was 6.86 cm^3 (95% CI: 5.48–8.24 cm^3), and from those using a semiautomated method was 4.63 cm^3 (95% CI: 1.54–7.72 cm^3), from those using an automated method was 4.47 cm^3 (95% CI: 3.68–5.26 cm^3) (Supplemental Fig. S6). Mean ages of healthy subjects in the only two studies that relied on manual assessment methods were 69.3 years[42] and 84.1 years.[49] Median ages and IQR of participants from studies included in the meta-analysis that relied on semiautomated and automated methods for the WMH volume assessment were 70.28 (IQR 55.79–72.08) years and 66.5 (IQR 62.02–70.63) years, respectively.

Discussion

To the extent of our knowledge, several studies[58–60] investigated the reproducibility of WMH volumetric measures within specific centers, but no report investigated intercenter

reproducibility.[11] The main result of our systematic review was a high heterogeneity in WMH volume. Even though we only selected studies with healthy subjects, risk factors for cerebral SVD, such as hypertension, low cardiac ejection fraction, carotid atherosclerosis, and atrial fibrillation, may have been unreported. Varying prevalence of these conditions may have contributed to the observed heterogeneity.[61]

We focused on age as the most important unmodifiable risk factor for WMH volume.[6] When considering only subjects older than 50 years, the meta-regression analysis showed a positive significant correlation between WMH volume and age. Age explained less than one-third of the variance in WMH volume. Apart from potential unreported conditions, the remaining variance could be explained by a combination of the WMH biological (ie, intrinsic) and technical variability.

Since translational neuroimaging studies aim at extracting the potential diagnostic and prognostic value of imaging biomarkers,[62] it is important to reveal the WHM biological variability by removing or modeling the technical sources of bias in the WMH assessment process.

The majority of the analyzed study groups employed 3T magnets. The pooled WMH volume was lower in studies employing 3T scanners than in those using 1.5T scanners. This finding is partially in contrast with a previous study that reported an underestimation of WMH volume with lower field scanners, but also increased nonfocal WMH and flow artifacts at higher field strength.[63] Notably, the 3T studies in our meta-analysis included the three study groups with a mean age less than 50 years and a mean WMH volume close to 0 cm^3. Including these relatively young, low WMH volume groups in our analysis has contributed to the observed finding.

The level of automation in WMH segmentation was also significantly different across studies. Studies of relatively young subjects generally adopted automated methods, while WHM tended to be manually segmented in studies involving older participants. This difference in age, more than any technical aspects related to partial volume effect and thresholding, may have contributed to the observed differences in the WMH pooled volumes.

Optimally, to keep technical variability to a minimum, standardized imaging protocols for investigating SVD should be used.[3] It is advisable to deploy high-resolution isotropic FLAIR imaging on 3.0T scanners to study WMH.[3] Harmonization strategies such as correction for B_1 inhomogeneities should then be adopted to reduce heterogeneity within and across MRI scanners.[64] Moreover, the availability of several automated methods for WMH segmentation may represent a technical source of bias.[10,65,66] The best-performing method should be identified among coexisting solutions after extensive testing on a common external dataset segmented by multiple experts and combined by consensus or using label fusion algorithms.[67]

Rigorous anatomical definition of WMH is also of paramount importance. Previous initiatives for standardization and harmonization in WMH assessment did not address the sources of variability in the segmentation process.[13] For instance, the innermost segment of periventricular WMH could either be segmented as WMH[68] or disregarded as partial volume or cerebrospinal fluid flow artifact.[69] Moreover, thin white matter bundles such as structures that belong to the limbic system and white matter bundles intermixed with deep gray matter structures are usually not considered in the anatomical operational definition of white matter used in structural brain MRI segmentation, although WMH in these bundles may carry significant clinical impact.[70] Pathology and clinically oriented study should guide the choice towards a standardized criterion for WMH segmentation. An endorsed anatomical definition of WMH would also increase the performance of supervised segmentation algorithms by ameliorating the quality of the labeled training set.

Limitations

The main limitation of this study was the relatively small number of meta-analyzed study groups, which, combined with a high heterogeneity among them, impeded us from providing a point estimate and reference normality interval for WMH volume in healthy adults. To increase the availability of studies, it is essential that descriptive statistics of the original imaging data are always reported alongside the transformed metrics.

Moreover, in this review we did not address the issue of slice thickness. The impact of slice thickness on lesion detection has been investigated in multiple sclerosis. In these studies, authors found only minor differences in the distribution of texture analysis parameter values in white matter lesions for 1-mm and simulated 3-mm-thick slices.[71] In our review and meta-analysis, slice thickness values were highly heterogeneous across the included studies. Most recent studies, especially those preformed at 3T, employed volumetric acquisitions with about 1 mm isotropic voxel size.[21,23,26,28,57] Other studies acquired axial slices with 2,[27,52] 3,[19,42,46] 4,[48,49] or even 5 mm[31,39] slice thickness. Thicker slices can lead to both overestimation of larger lesions and underestimation of smaller lesions. For this reason, we cannot exclude that varying slice thickness across the examined studies might have contributed to the observed differences in pooled WMH volume between scanner strength and level of automation subgroups.

Conclusion

At the current time, data on WMH volume in healthy adults appear to be not comparable across studies. We stress the need for achieving better standardization in WMH

quantification and reporting and encourage international initiatives aimed at promoting this.

Acknowledgments

This study was partially supported by Ricerca Corrente funding from the Italian Ministry of Health to IRCCS Policlinico San Donato. This study received no specific grant from any funding agency in the public, commercial, or nonprofit sectors.

References

1. Wardlaw JM, Valdés Hernández MC, Muñoz-Maniega S. What are white matter hyperintensities made of? Relevance to vascular cognitive impairment. J Am Heart Assoc 2015;4(6):e001140.
2. Debette S, Schilling S, Duperron MG, Larsson SC, Markus HS. Clinical significance of magnetic resonance imaging markers of vascular brain injury: A systematic review and meta-analysis. JAMA Neurol 2019;76(1):81-94.
3. Wardlaw JM, Smith EE, Biessels GJ, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol 2013;12(8):822-838.
4. Sarbu N, Shih RY, Oleaga L, Smirniotopoulos JG. Radiographics update: White matter diseases with radiologic-pathologic correlation. Radiographics 2020;40(3):E4-E7.
5. Alber J, Alladi S, Bae H, et al. White matter hyperintensities in vascular contributions to cognitive impairment and dementia (VCID): Knowledge gaps and opportunities. Alzheimers Dement (NY) 2019;5(1):107-117.
6. Grueter BE, Schulz UG. Age-related cerebral white matter disease (Leukoaraiosis): A review. Postgrad Med J 2012;88(1036):79-87.
7. Frey BM, Petersen M, Mayer C, Schulz M, Cheng B, Thomalla G. Characterization of white matter hyperintensities in large-scale MRI-studies. Front Neurol 2019;10:238.
8. Van den Heuvel DMJ, Ten Dam VH, De Craen AJM, et al. Measuring longitudinal white matter changes: Comparison of a visual rating scale with a volumetric measurement. Am J Neuroradiol 2006;27(4):875-878.
9. Olsson E, Klasson N, Berge J, et al. White matter lesion assessment in patients with cognitive impairment and healthy controls: Reliability comparisons between visual rating, a manual, and an automatic volumetrical MRI method—The Gothenburg MCI study. J Aging Res 2013;2013:1-10.
10. Caligiuri ME, Perrotta P, Augimeri A, Rocca F, Quattrone A, Cherubini A. Automatic detection of white matter hyperintensities in healthy aging and pathology using magnetic resonance imaging: A review. Neuroinformatics 2015;13(3):261-276.
11. De Guio F, Jouvent E, Biessels GJ, et al. Reproducibility and variability of quantitative magnetic resonance imaging markers in cerebral small vessel disease. J Cereb Blood Flow Metab 2016;36(8):1319-1337.
12. Woo CW, Chang LJ, Lindquist MA, Wager TD. Building better biomarkers: Brain models in translational neuroimaging. Nat Neurosci 2017;20(3):365-377.
13. Smith EE, Biessels GJ, De Guio F, et al. Harmonizing brain magnetic resonance imaging methods for vascular contributions to neurodegeneration. Alzheimers Dement (Amst) 2019;11:191-204.
14. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med 2009;6(7):e1000097.
15. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 2010;1(2):97-111.
16. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21(11):1539-1558.
17. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-188.
18. Sardanelli F, Di Leo G. *Biostatistics for radiologists: Planning, performing, and writing a radiologic study*. Berlin: Springer; 2009.
19. Huang CC, Yang AC, Chou KH, et al. Nonlinear pattern of the emergence of white matter hyperintensity in healthy Han Chinese: An adult lifespan study. Neurobiol Aging 2018;67:99-107.
20. Guzman-Ortiz E, Bueno-Hernandez N, Melendez-Mier G, Roldan-Valadez E. Quantitative systematic review: Methods used for the in vivo measurement of body composition in pregnancy. J Adv Nurs 2020;1–13. https://doi.org/10.1111/jan.14594 [epub ahead of print].
21. Jäncke L, Liem F, Merillat S. Weak correlations between body height and several brain metrics in healthy elderly subjects. Eur J Neurosci 2019;50(10):3578-3589.
22. Damulina A, Pirpamer L, Seiler S, et al. White matter hyperintensities in Alzheimer's disease: A lesion probability mapping study. J Alzheimers Dis 2019;68(2):789-796.
23. Lampe L, Kharabian-Masouleh S, Kynast J, et al. Lesion location matters: The relationships between white matter hyperintensities on cognition in the healthy elderly. J Cereb Blood Flow Metab 2019;39(1):36-43.
24. Lau AY, Mok V, Lee J, et al. Retinal image analytics detects white matter hyperintensities in healthy adults. Ann Clin Transl Neurol 2019;6(1):98-105.
25. Seiler S, Fletcher E, Hassan-Ali K, et al. Cerebral tract integrity relates to white matter hyperintensities, cortex volume, and cognition. Neurobiol Aging 2018;72:14-22.
26. Nylander R, Kilander L, Ahlström H, Lind L, Larsson E-M. Small vessel disease on neuroimaging in a 75-year-old cohort (PIVUS): Comparison with cognitive and executive tests. Front Aging Neurosci 2018;10:217.
27. Phyu P, Merwick A, Davagnanam I, et al. Increased resting cerebral blood flow in adult Fabry disease: MRI arterial spin labeling study. Neurology 2018;90(16):E1379-E1385.
28. Van Rooden S, Van Den Berg-Huysmans AA, Croll PH, et al. Subjective cognitive decline is associated with greater white matter hyperintensity volume. J Alzheimers Dis 2018;66(3):1283-1294.
29. Vipin A, Foo HJL, Lim JKW, et al. Regional white matter hyperintensity influences grey matter atrophy in mild cognitive impairment. J Alzheimers Dis 2018;66(2):533-549.
30. Weinstein G, Zelber-Sagi S, Preis SR, et al. Association of nonalcoholic fatty liver disease with lower brain volume in healthy middle-aged adults in the Framingham Study. JAMA Neurol 2018;75(1):97-104.
31. Ye Q, Su F, Gong L, et al. Divergent roles of vascular burden and neurodegeneration in the cognitive decline of geriatric depression patients and mild cognitive impairment patients. Front Aging Neurosci 2017;9:288.
32. Moura AR, Lee S, Habeck C, Razlighi Q, Stern Y. The relationship between white matter hyperintensities and cognitive reference abilities across the life span. Neurobiol Aging 2019;83:31-41.
33. De Marco M, Manca R, Mitolo M, Venneri A. White matter hyperintensity load modulates brain morphometry and brain connectivity in healthy adults: A neuroplastic mechanism? Neural Plast 2017;2017:1-10.
34. Devantier TA, Nørgaard BL, Poulsen MK, et al. White matter lesions, carotid and coronary atherosclerosis in late-onset depression and healthy controls. Psychosomatics 2016;57(4):369-377.
35. Liu J, Tseng BY, Khan MA, et al. Individual variability of cerebral autoregulation, posterior cerebral circulation and white matter hyperintensity. J Physiol 2016;594(11):3141-3155.
36. Lockhart SN, Luck SJ, Geng J, et al. White matter hyperintensities among older adults are associated with futile increase in frontal activation and functional connectivity during spatial search. PLoS One 2015;10(3):e0122445.

37. Jiang J, Trollor JN, Brown DA, et al. An inverse relationship between serum macrophage inhibitory cytokine-1 levels and brain white matter integrity in community-dwelling older individuals. Psychoneuroendocrinology 2015;62: 80-88.

38. Reijmer YD, Schultz AP, Leemans A, et al. Decoupling of structural and functional brain connectivity in older adults with white matter hyperintensities. Neuroimage 2015;117:222-229.

39. Ternifi R, Cazals X, Desmidt T, et al. Ultrasound measurements of brain tissue pulsatility correlate with the volume of MRI white-matter hyperintensity. J Cereb Blood Flow Metab 2014;34(6):942-944.

40. Tarumi T, Ayaz Khan M, Liu J, et al. Cerebral hemodynamics in normal aging: Central artery stiffness, wave reflection, and pressure pulsatility. J Cereb Blood Flow Metab 2014;34(6):971-978.

41. Huang CC, Liu ME, Chou KH, et al. Effect of BDNF Val66Met polymorphism on regional white matter hyperintensities and cognitive function in elderly males without dementia. Psychoneuroendocrinology 2014;39 (1):94-103.

42. Rowe Bijanki K, Arndt S, Magnotta VA, et al. Characterizing white matter health and organization in atherosclerotic vascular disease: A diffusion tensor imaging study. Psychiatry Res 2013;214(3):389-394.

43. Yatawara C, Lee D, Ng KP, et al. Mechanisms linking white matter lesions, tract integrity, and depression in Alzheimer disease. Am J Geriatr Psychiatry 2019;27(9):948-959.

44. Birdsill AC, Carlsson CM, Willette AA, et al. Low cerebral blood flow is associated with lower memory function in metabolic syndrome. Obesity 2013;21(7):1313-1320.

45. Gurol ME, Viswanathan A, Gidicsin C, et al. Cerebral amyloid angiopathy burden associated with leukoaraiosis: A positron emission tomography/magnetic resonance imaging study. Ann Neurol 2013;73 (4):529-536.

46. Makedonov I, Black SE, Macintosh BJ. Cerebral small vessel disease in aging and Alzheimer's disease: A comparative study using MRI and SPECT. Eur J Neurol 2013;20(2):243-250.

47. He J, Carmichael O, Fletcher E, et al. Influence of functional connectivity and structural MRI measures on episodic memory. Neurobiol Aging 2012;33(11):2612-2620.

48. Portet F, Brickman AM, Stern Y, et al. Metabolic syndrome and localization of white matter hyperintensities in the elderly population. Alzheimers Dement 2012;8(Suppl 5):S88-S95.e1.

49. Silbert LC, Howieson DB, Dodge H, Kaye JA. Cognitive impairment risk: White matter hyperintensity progression matters. Neurology 2009; 73(2):120-125.

50. Koga H, Takashima Y, Murakawa R, Uchino A, Yuzuriha T, Yao H. Cognitive consequences of multiple lacunes and leukoaraiosis as vascular cognitive impairment in community-dwelling elderly individuals. J Stroke Cerebrovasc Dis 2009;18(1):32-37.

51. Holland CM, Smith EE, Csapo I, et al. Spatial distribution of white-matter hyperintensities in Alzheimer disease, cerebral amyloid angiopathy, and healthy aging. Stroke 2008;39(4):1127-1133.

52. de Schipper LJ, Hafkemeijer A, Bouts MJRJ, et al. Age- and disease-related cerebral white matter changes in patients with Parkinson's disease. Neurobiol Aging 2019;80:203-209.

53. Bronas UG, Steffen A, Dlon C, et al. Sedentary time and white matter hyperintensity volume in older adults. Med Sci Sports Exerc 2019;51(8): 1613-1618.

54. Nam KW, Kwon HM, Jeong HY, Park JH, Kwon H, Jeong SM. Serum homocysteine level is related to cerebral small vessel disease in a healthy population. Neurology 2019;92(4):E317-E325.

55. Nam K-W, Kwon H-M, Jeong H-Y, Park J-H, Kwon H, Jeong S-M. Cerebral small vessel disease and stage 1 hypertension defined by the 2017 American College of Cardiology/American Heart Association guidelines. Hypertension 2019;73(6):1210-1216.

56. Zhang CE, Wong SM, Uiterwijk R, et al. Blood-brain barrier leakage in relation to white matter hyperintensity volume and cognition in small vessel disease and normal aging. Brain Imaging Behav 2019;13(2): 389-395.

57. Ye Q, Chen X, Qin R, et al. Enhanced regional homogeneity and functional connectivity in subjects with white matter hyperintensities and cognitive impairment. Front Neurosci 2019;13:695.

58. de Boer R, Vrooman HA, Ikram MA, et al. Accuracy and reproducibility study of automatic MRI brain tissue segmentation methods. Neuroimage 2010;51(3):1047-1056.

59. Wei X, Warfield SK, Zou KH, et al. Quantitative analysis of MRI signal abnormalities of brain white matter with high reproducibility and accuracy. J Magn Reson Imaging 2002;15(2):203-209.

60. Ramirez J, Scott CJM, Black SE. A short-term scan-rescan reliability test measuring brain tissue and subcortical hyperintensity volumetrics obtained using the lesion explorer structural MRI processing pipeline. Brain Topogr 2013;26(1):35-38.

61. Moroni F, Ammirati E, Rocca MA, Filippi M, Magnoni M, Camici PG. Cardiovascular disease and brain health: Focus on white matter hyperintensities. Int J Cardiol Heart Vasc 2018;19:63-69.

62. Chertkow H, Black S. Imaging biomarkers and their role in dementia clinical trials. Can J Neurol Sci 2007;34(Suppl 1):S77-S83.

63. Neema M, Guss ZD, Stankiewicz JM, Arora A, Healy BC, Bakshi R. Normal findings on brain fluid-attenuated inversion recovery MR images at 3T. Am J Neuroradiol 2009;30(5):911-916.

64. Guo C, Niu K, Luo Y, et al. A comprehensive study on reproducibility of automatic white matter hyperintensities quantification. Front Neurosci 2019;13(JUL):679.

65. Heinen R, Steenwijk MD, Barkhof F, et al. Performance of five automated white matter hyperintensity segmentation methods in a multicenter dataset. Sci Rep 2019;9(1):16742.

66. Steenwijk MD, Pouwels PJW, Daams M, et al. Accurate white matter lesion segmentation by k nearest neighbor classification with tissue type priors (kNN-TTPs). Neuroimage Clin 2013;3:462-469.

67. Gibson E, Hu Y, Huisman HJ, Barratt DC. Designing image segmentation studies: Statistical power, sample size and reference standard quality. Med Image Anal 2017;42:44-59.

68. Valdés Hernández MC, Piper RJ, Bastin ME, et al. Morphologic, distributional, volumetric, and intensity characterization of periventricular hyperintensities. Am J Neuroradiol 2014;35(1):55-62.

69. Payne ME, Fetzer DL, MacFall JR, Provenzale JM, Byrum CE, Krishnan KRR. Development of a semiautomated method for quantification of MRI gray and white matter lesions in geriatric subjects. Psychiatry Res 2002;115(1–2):63-77.

70. Vangberg TR, Eikenes L, Håberg AK. The effect of white matter hyperintensities on regional brain volumes and white matter microstructure, a population-based study in HUNT. Neuroimage 2019;203:116158.

71. Savio SJ, Harrison LCV, Luukkaala T, et al. Effect of slice thickness on brain magnetic resonance image texture analysis. Biomed Eng Online 2010;9:60.